University of Alberta

The Test Retest Reliability, Construct Validity, and Responsiveness of the Tennis Elbow Function Scale

by

Katherine Audrey Lowe

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment

of

the requirements for the degree of Master of Science

Department of Physical Therapy

Edmonton, Alberta
Fall 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-47060-1

Canada

University of Alberta

Library Release Form

Name of Author: *Katherine Audrey Lowe*

Title of Thesis: *The Test Retest Reliability, Construct Validity, and Responsiveness of the Tennis Elbow Function Scale*

Degree: *Master of Science*

Year this Degree Granted: *1999*

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Katherine Audrey Lowe

9515 77 Street
Edmonton, Alberta
T6C 2M8

August 24, 1999

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *The Test Retest Reliability, Construct Validity and Responsiveness of the Tennis Elbow Function Scale* submitted by *Katherine Audrey Lowe* in partial fulfillment of the requirements for the degree of Master of Science

Michele Cotes-Battie

Jean Wessel

Sharon Warren

Robert Burnham

July 19, 1999

DEDICATION

This work is dedicated

to the memory of Warren Lowe,

who taught me that the greatest deeds are those that involve helping others,

to my parents,

Dolly Lowe and Evan Lowe,

who taught me to persevere in the face of adversity

and to

Terri Kaasa, Kari Elliot, Connie Winther and Jack Seto,

who taught me to laugh in spite of adversity.

Abstract

This study examined the reliability, construct validity and responsiveness of the Tennis Elbow Function Scale (TEFS), a measure of pain-related elbow function. Using a single-group repeated measures study design, thirty-one volunteers with lateral epicondylitis were tested three times during a seven-week period. The TEFS and Pain Free Function Index (PFF) were administered at all three tests. The visual analogue scales for pain (pain VAS) and function (function VAS) and pain free grip strength (PFG) were administered at Test 1 and Test 3. At Test 3, therapists and subject rated changes in their condition on change rating scales (CRS).

The ICC 1,1 was 0.92 (95% CI 0.83- 0.96). The TEFS was moderately correlated with the other measures and it was more efficient than the PFF in evaluating improved subjects. The results indicate that the TEFS is a reliable, valid and responsive measure. It should be considered suitable for evaluating patients with LE.

ACKNOWLEDGEMENTS

This project would never have been realized without the support from two special people:

- my supervisor, Dr. Michele Crites-Battié, who spent never ending hours editing this project and setting goals and deadlines which I never kept;
- my interim supervisor, Dr. Jean Wessel, who taught me many things about measurement and evaluation and who helped me organize disorganized thoughts into something readable.

It has been an honor and privilege to work with both of these dedicated researchers and teachers, and to observe and benefit from their talents.

I would also like to thank:

- my committee members, Dr. Sharon Warren and Dr. Robert Burnham,
- the developer of the Tennis Elbow Function Scale, Mr. Paul Stratford

Financial support was received from the Capital Health Authority – Orthopaedic Research Fund, the Health Sciences Association – George C. Hall Bursary, Strathcona Physiotherapy Clinic – Physiotherapy Research Award, and the University of Alberta - Walter H. Johns Graduate Tuition Scholarship.

Table of Contents

INTRODUCTION 1

Problem Statement 1

PURPOSE OF THE STUDY 3

Objectives 3
Limitations 5
Delimitations 5
Definitions 6
Abbreviations 7

LITERATURE REVIEW 8

Epidemiology of Lateral Epicondylitis 8
Research of Treatment Efficacy for Lateral Epicondylitis 9
Properties of Evaluative Health Measures 11
Outcome Measures in Lateral Epicondylitis Clinical Trials 13

METHODS 17

Study Design 17
Sample Size Considerations 17
Subject Recruitment 18
Outcome Measures 19
Data Collection 21
Analysis 22
Ethical Considerations 23

RESULTS 25

Subject Characteristics 25
Analysis of the TEFS 26

DISCUSSION 33

Generalizability of Results 33
Reliability 34
Construct Validity 36
Responsiveness 39
General Observations 42
Suggestions for Future Research 45

CONCLUSION 46

REFERENCES 48

APPENDICES 56

Appendix A Patient-Rated Tennis Elbow Evaluation 57
Appendix B Tennis Elbow Function Scale 58
Appendix C Sample Size Calculation Intraclass Correlation 59
Appendix D Master Table for Sample Size Calculation 60
Appendix E Sample Size for Pearson Product-Moment Correlation 61
Appendix F Physical Therapy Clinics 62
Appendix G Inclusion Criteria Screen 63

Appendix H	Consent Form	64
Appendix I	Patient Information Letter	65
Appendix J	Visual Analogue Scales for Pain and Function	67
Appendix K	Subject Change Rating Scale	68
Appendix L	Therapist Change Rating Scale	69
Appendix M	Reliability Analysis	70
Appendix N	Responsiveness Analysis	71
Appendix O	Letter of Approval from the Health Research Ethics Board	72
Appendix P	Letter of Approval from Caritas Research Steering Committee	73

List of Tables

TABLE 1: DESCRIPTIVE STATISTICS FOR OUTCOMES MEASURES TAKEN ON THE THREE TEST SESSIONS FOR ALL SUBJECTS (N=31) 27
TABLE 2: DESCRIPTIVE STATISTICS FOR THE DIFFERENCE SCORES OF OUTCOME MEASURES FOR ALL SUBJECTS (N=31) 27
TABLE 3: PEARSON CORRELATION COEFFICIENTS AND P-VALUES () FOR MEASURES AT TEST 1 (N=31) 28
TABLE 4: COMPARISON OF TEFS SCORES AT TEST 1,2, AND 3 28
TABLE 5: RESPONSIVENESS INDEX FOR ALL SUBJECTS AND IMPROVED SUBJECTS 29
TABLE 6: PEARSON CORRELATION COEFFICIENTS AND P-VALUES () FOR CHANGE SCORES (N=25) 29

List of Figures

FIGURE 1: DISTRIBUTION OF TEST 1 TEFS SCORES 30
FIGURE 2: DISTRIBUTION OF TEST 1 PFF SCORES 30
FIGURE 3: CORRELATION OF THE TEST 1 TEFS AND PFF 31
FIGURE 4: CORRELATION OF THE TEST 1 TEFS AND TEST 1 FUNCTION VAS 31

Chapter 1

Introduction

Problem Statement

Physical therapists require outcome measures which can be used to evaluate the function of patients with lateral epicondylitis (LE), a musculoskeletal condition characterized by tenderness over the lateral epicondyle and by lateral forearm pain increased with activity. LE occurs in one to three percent of the general adult population.[1] As many as 40 to 47% of tennis players will develop LE.[2] Traditionally LE has been treated using a variety of modalities, but investigation of various treatments provides questionable evidence of their efficacy.[3] The trials investigating LE treatments are characterized by a lack of controls, a lack of randomization, and limited use of validated outcome measures.[3, 4] Analyses of 78 LE treatment studies[3] conducted prior to 1990 found that only one trial had a study design with sufficient internal controls to qualify it as a valid clinical trial. Subsequent studies[5-8] have improved in the areas of randomization, follow-up periods and appropriate statistical analysis. However, most studies have not used valid measures of function.

Only two clinical trials have used a validated outcome measure of function for LE.[7, 9] Both trials used the Painfree Function Index (PFF), an eight-item scale requiring clients to indicate if elbow discomfort is present or absent during the performance of activities requiring forearm movement or gripping (Appendix A). A limitation of the

PFF is the dichotomous response scale, which is not well suited to evaluating clinical change. Another version of the scale has been developed. The Tennis Elbow Function Scale (TEFS) is a ten-item, self-report scale designed to measure elbow discomfort during the performance of personal care, household, work and recreational activities (Appendix B). The TEFS has a five-point response scale, therefore it has the potential to be more responsive than the PFF for evaluating clinical change. The TEFS has not been used in any published studies. Research is required to examine the measurement properties of the TEFS and to determine if it is suitable for use in clinical trials for LE.

Purpose of the Study

The purpose of the study was to evaluate the reliability, construct validity and responsiveness of the TEFS.

Objectives

1. Determine the test retest reliability of the TEFS in subjects with LE.

Hypothesis 1: The intraclass correlation coefficient of TEFS scores measured at Test 1 and Test 2 is high (ICC [1,1] > 0.75).

2. Examine the construct validity of the TEFS in subjects with LE.

Hypothesis 2: The TEFS is considered valid if Test 1 scores of the TEFS are:

a) moderately positively correlated with Test 1 scores for pain and function as measured by the visual analogue scales for pain (pain VAS) and function (function VAS) ($r > 0.50$).

b) moderately negatively correlated with Test 1 pain free grip strength (PFG) ($r > -0.50$)

c) strongly correlated with Test 1 PFF scores ($r > 0.75$).

d) not significantly correlated with maximum grip strength scores on the unaffected arm (MGU).

3) Examine the responsiveness of the TEFS in subjects with LE.

Hypothesis 3: Scores of TEFS are lower at Test 3 than at Test 2 and Test 1.

Hypothesis 4: The Relative Efficiency (RE) of the TEFS compared to the PFF is >1 in subjects rated as improved.

Hypothesis 5: The Responsiveness Index (RI) is greater for the TEFS compared to the PFF.

Hypothesis 6: In subjects rated as improved, the change between Test 1 and Test 3 scores of the TEFS is:

a) moderately positively correlated ($r > 0.50$) with changes between Test 1 and Test 3 scores of the pain VAS, function VAS and PFF.

b) moderately negatively correlated ($r > -0.50$) with changes between Test 1 and Test 3 PFG scores.

Limitations

- The reliability and validity of the Change Rating Scale used by both the therapist and subject has not been established.
- The effectiveness of the treatments used during the interval between test sessions is not known but the treatments and the passage of time are presumed to effect an improvement in the condition.

Delimitations

- The results of the reliability, construct validity and responsiveness analysis for the TEFS are applicable to those people with a clinical diagnosis of lateral epicondylitis who seek physical therapy treatment and do not have the comorbid conditions listed in the inclusion criteria.

Definitions

Lateral Epicondylitis

a musculoskeletal condition characterized by tenderness over the lateral epicondyle, and lateral forearm pain that increases with activity.[10]

Reliability

The degree of consistency with which an instrument or rater measures a variable.[11]

Construct Validity

A type of measurement validity, the degree to which a theoretical construct is measured by an instrument.[11]

Responsiveness

An instrument's ability to detect clinically significant change.[12]

Abbreviations

ANOVA	Analysis of Variance
Subject CRS	Subject Change Rating Scale
Therapist CRS	Therapist Change Rating Scale
ICC	Intraclass Correlation Coefficient
LE	Lateral Epicondylitis
PFF	Painfree Function Index
PFG	Pain free Grip Strength
MGU	Maximum Grip Strength of the Unaffected Arm
RE	Relative Efficiency
RI	Responsiveness Index
SEM	Standard Error of Measurement
TEFS	Tennis Elbow Function Scale
VAS	Visual Analogue Scale

Chapter 2

Literature Review

Epidemiology of Lateral Epicondylitis

LE is a musculoskeletal condition that affects the tendinous origin of the wrist extensor musculature.[10] It is primarily characterized by pain in the lateral aspect of the elbow and forearm.[10, 13] The symptoms are exacerbated with use of the forearm or hand musculature.[10] Generally, people with LE seek medical treatment one to twelve months following the onset of symptoms.[14] The diagnosis is made primarily from a clinical examination, which reveals tenderness on palpation at the lateral epicondyle and pain with resisted wrist or finger extension.[10, 15] Decreased grip strength, limitation of elbow and wrist movement, and functional limitations due to pain are also commonly reported.[15-19]

LE was first described in the late 1870s and was associated with tennis playing.[20] Recreational or occupational activities requiring prolonged use of the forearm musculature or repetitive gripping may predispose one to acquiring this condition.[10] It is prevalent in one to three percent in the general adult population.[1, 6] As many as 40 to 47% of tennis players will develop LE.[2, 15] Older tennis players who practice more frequently for longer periods are most susceptible to developing LE.[15]

In industry the prevalence is 0.8% in workers in the meat packing industry and 2 to 3% in female shop and assembly line workers.[21, 22] Jobs with high repetition rates

and forceful wrist movements are cited as contributing to the development of the condition.[23]

Cyriax[24] claimed LE was a self-limiting condition. However, this assumption may not be accurate. For example, Binder[14] reported that 50% of people treated one to five years previously continued to experience symptoms which required them to modify their activities, but not seek further medical treatment. Thus, patients may continue to experience symptoms long after being discharged from medical treatment. Also recurrence of symptoms has been reported in 3%[25] to 54%[26] of cases, and the recurrences may be underreported because people stop seeking medical treatment.[14, 24, 27]

For those patients with symptoms of shorter duration, there may be a better prognosis.[4-6] In subjects with symptoms lasting less than three months, steroid injections resulted in greater pain reduction.[5] Subjects with symptoms lasting less than six months were more likely to completely recover following the administration of various physical therapy treatments.[2]

Research of Treatment Efficacy for Lateral Epicondylitis

Clinical trials investigating the efficacy of cortisone injections, transverse frictions, ultrasound, electrotherapy, short wave diathermy, acupuncture and bracing have failed to conclusively demonstrate the superiority of a particular treatment.[17, 28-35] For example, questions arise concerning the effects of ultrasound in the treatment of LE. Continuous ultrasound treatments were considered to produce better results compared to

rest, but produced equivocal results compared to placebo ultrasound.[29] Although ultrasound was noted to be less effective than steroid injections in relieving pain immediately after treatments were completed,[14] six-month recurrence rates were lower in subjects treated with ultrasound alone and in subjects treated with ultrasound prior to receiving cortisone injections.[14]

Failure to demonstrate efficacy in some studies may accurately reflect the administration of ineffective treatment.[3] LE is characterized by degenerative rather than inflammatory changes at the tendon of extensor carpi radialis brevis.[13, 14] Modalities such as steroid injections are directed towards controlling local sources of inflammation and pain. Thus the steroid injections are directed towards treating symptoms rather than treating the underlying pathology of LE.[14, 28]

Poor study design is another explanation for failure to demonstrate the efficacy of a particular treatment for LE. Published research on treatment efficacy is characterized by a lack of randomization and controls.[3] For example, Labelle[3] searched the Medline and EMbase databases for all published clinical treatment trials conducted between 1960 and 1990, and found that only eighteen of 78 clinical trials were controlled. The technique of Chalmers[36] was used to analyze the quality of study design and data analyses. Of the eighteen clinical trials examined, only one study was considered acceptable as a valid clinical trial. Based upon the lack of valid clinical trials for LE, Labelle et al.[3] concluded that there was insufficient scientific evidence to support most of the current treatment methods. Studies examining the role of exercise or counterforce bracing published after 1990 have improved in the areas of randomization, follow-up

period and data analysis.[5-8] However, crossover study designs, inadequate blinding of raters, subject drop out, and migration to other treatments are problems with recent research.[5, 37]

Finally, the measurement tools used to evaluate outcomes may be responsible for failure to demonstrate conclusively the efficacy of a particular treatment. Only two studies have reported the properties of reliability, validity and responsiveness of all outcome measures used to monitor changes in LE.[7, 17] Because most clinical trials do not report the psychometric properties of the outcome measures used,[28, 29, 33-35, 38-41] the results may not accurately reflect changes that may have occurred.

Properties of Evaluative Health Measures

Kirshner and Guyatt[42] recommend that health indices used for evaluative purposes undergo investigation to determine their reliability, validity and responsiveness. Reliability is a measure of the degree of consistency with which an instrument measures a stable variable.[11] Outcome measures that demonstrate little variability in scores measured under stable conditions on repeated occasions are considered reliable.[42] It is important to use reliable outcome measures because they provide a degree of confidence that changes in the outcome measure can be attributed to changes in the condition rather than error.

Reliability is a necessary but not a sufficient condition for an evaluative health measure.[42] For example, maximum grip strength in someone with elbow pain is considered reliable, but it may not be the most valid method of detecting improvement in

LE subjects. Therefore, maximum grip strength may not be the most suitable measure for evaluating changes in lateral elbow pain.[43]

Validity refers to the degree to which the instrument measures what it is intended to measure.[11, 44] One type of validity, construct, is often evaluated by examining the relationships between the measure of interest and instruments that measure similar or different constructs.[11, 12] If the correlations are of the expected magnitude and direction, the construct validity of the instrument is supported.[12] For example, the construct validity of region-specific disability scales is often determined by examining the relationship with generic health-status scales and practitioners' rating of function.[45, 46]

Responsiveness refers to an instrument's ability to detect clinically significant changes.[12, 47] One method to evaluate responsiveness is to measure the effect size, that is the magnitude of change following treatment.[48, 49] The relative efficiency index (RE) can be used to compare the effect size of two different outcome measures administered to the same group of patients whose health status has changed.[50] A ratio is calculated between the paired t-scores of the two scales.[50] The measure with the greatest t-score is designated as being the most efficient at detecting change.

Change scores can be examined relative to variability in subjects whose condition is stable. The responsiveness index (RI) is a ratio derived from the variability of test scores in stable subjects and change scores in subjects who have undergone treatment and are expected to have changes.[47] This analysis requires that the reliability of the measure be known. Variability of test scores in stable subjects can be determined in a separate

study or alternatively by using a repeated-measures design in which multiple baseline measures are administered to subjects prior to and following a treatment intervention.

Finally, change scores of a newly developed instrument can be validated through correlation with other impairment or health measures monitoring similar or different dimensions of a clinical condition. For example, changes in disease severity measures such as joint tenderness count, and cardiac or respiratory status have been used to validate newly developed self-report functional status scales.[42, 51] When the change scores of each measure are related in a manner as theoretically predicted, then the functional status measure is also considered responsive to changes in the clinical condition.[12]

Outcome Measures in Lateral Epicondylitis Clinical Trials

The outcomes most commonly assessed in LE clinical trials are pain, strength and function. Only a few measures have been validated for use in this population. The pain measure most frequently used in clinical trials is the visual analogue scale (pain VAS).[17, 31, 52] It is a ten-centimetre line with its ends defined to represent the maximum and minimum extremes of pain intensity.[53] Subjects mark the line to indicate the severity of their pain. In LE subjects, the reliability of a vertical pain VAS has been reported to be strong ($r = 0.89$).[9] It is considered a valid measure because it has been correlated with pain free grip strength ($r = -0.47$) and a pain-free function measure ($r = 0.68$).[9] The pain VAS is considered a responsive outcome measure for LE because mean pain VAS change scores were greater in subjects who successfully responded to treatment of LE

than in those whose treatment was considered a failure (success 23.2 mm vs. failure -2.5 mm).[54]

The reliable grip measures that have been used in clinical trials of LE are pain-free grip (PFG) and maximum grip strength (MG). Both measures are considered valid for this population because they have been related to measures of pain and function. PFG is moderately to strongly related to unidimensional pain and function measures (r = - 0.66 for a pain VAS[9] and 0.98 for a numeric rating scale[55] and r = - 0.47 for a function VAS[9] and 0.99 for a numeric rating scale).[55] Compared to maximum grip, PFG demonstrates stronger correlations with overall pain and function measures.[9] It is more responsive than maximum grip in evaluating changes in those responding successfully to treatment.[54, 56]

Isokinetic forearm and wrist strength of LE subjects have been found to change following surgery or exercise therapy.[31, 41] Pienimaki[31] advocated that isokinetic strength measures are important measures of arm function. However, the relationship between isokinetic strength and arm function was not reported by the researcher.[31]

There is no criterion measure for elbow function. Some work has been conducted to develop single-item, self-report measures of overall function. Burton[55] first investigated the construct of overall function in subjects with LE by using a six-point numeric rating scale. In this study estimations of overall function were strongly correlated with estimations of overall pain and PFG (r = - 0.98 and 0.99).[55] Another study found only moderate correlations between visual analogue scales for function (r = - 0.49) and pain (r = 0.66) and PFG (r = 0.66).[9] The function VAS was also found to be

reliable.[9] Single-item scales do not, however, detail the specific functions or activities limited by LE.

Pienimaki[31] used a pain and disability questionnaire in a clinical trial of LE. Subjects were asked to estimate sleep disturbances, pain at rest and under strain, and ability to work, lift objects, and perform hobbies.[31] The researcher did not provide information concerning the psychometric properties of this scale.[31]

Albrecht[41] modified a low back pain questionnaire used by the Japanese Orthopaedic Association for a study investigating the results of surgery for LE. This scale contains three sections rating subjective symptoms, clinical signs and impairment of functional activities. Scores for a control group were significantly higher than scores for LE patients. Overall, this scale appears to be responsive because scores changed following surgical intervention for LE. However, the important properties of construct validity and reliability have yet to be determined.

The PFF developed by Stratford et al.[9] is an eight-item, self-report measure of elbow discomfort experienced with the performance of specific activities.[9] The PFF was originally developed for a clinical trial evaluating ultrasound, phonophoresis and transverse frictions as treatments for LE.[17] Repeat measures taken within four days demonstrated good reliability (ICC = 0.93).[9] The PFF has been moderately correlated with overall function (function VAS, r = 0.70), PFG (r = 0.64) and pain (pain VAS, r = -0.68).[9] The PFF is considered a more efficient and responsive measure than the pain or

function VAS, and maximum grip strength when evaluating LE subjects responding successfully to treatment.[54]

Subsequent to the PFF, Stratford and Binkley[57] developed the Tennis Elbow Function Scale (TEFS). In this version, ten items were generated by LE subjects using the Patient Specific Functional Scale,[57, 58] a scale which asks individuals to list the activities that they are unable to do or have difficulty performing as a result of their problem.[58] An item analysis for the TEFS has been performed[57]. Preliminary analysis found the internal consistency of the TEFS to be 0.82.[57]

In the TEFS, eight questions deal with elbow discomfort experienced with the performance of gripping tasks such as dressing, opening a jar, writing and sweeping. Two questions pertain to the amount of discomfort experienced with the performance of usual work-related or recreational activities.

One of the limitations associated with the PFF is the dichotomous response scale. Such response scales are more appropriate for diagnostic rather than evaluative measures.[42] Using the PFF, subjects' elbow pain must be completely resolved before a test item score changes. Conversely, there is a five-point scale on the TEFS.[57] Subjects rate if an item causes no, slight, moderate, quite a bit, or extreme discomfort. Therefore, the TEFS has the potential to detect minimally clinically significant changes more efficiently than the PFF. Research is required to further test the psychometric properties of the TEFS.

Chapter 3

Methods

Study Design

This study used a single-group, repeated-measures design. On three occasions over a period of seven weeks, outcome measures were administered to subjects. Test 1 occurred within one week of the subject being confirmed as having LE at a screening visit and agreeing to participate. Test 2, occurred within four working days of Test 1. The final test occurred six weeks following Test 2, after treatment had been implemented.

Sample Size Considerations

A sample size of thirty was calculated using the formula and master table (Appendix C and D) from Kramer and Thiemann.[59] The calculation was based upon a hypothesized intraclass correlation of 0.75, a desired correlation of 0.90, an alpha level of 0.05 and a power level of 0.80. Fewer subjects were required for a sample size calculation using Pearson product-moment correlation (Appendix E).

It was anticipated that the desired sample size could be achieved in several months because two owners of the five participating clinics were consulted prior to commencing the study, and they estimated that in their respective clinics, three to five LE subjects were treated per month.[60]

Subject Recruitment

Subjects were volunteers receiving treatment for LE from eight physical therapy clinics (Appendix F). After the first three months, the subject recruitment at the five initial clinics was progressing slowly, so five other clinics were enlisted to act as referral sites. Only three of these clinics followed through with screening patients. The clinic owners reported that all LE clients receiving physical therapy treatment were screened for eligibility by their treating therapist on the first visit to the clinic (Appendix G).

Subjects were eligible for this study if they satisfied all of the following inclusion criteria:

1. complaints of lateral elbow pain for a minimum duration of one week
2. possession of two of the following four signs:
    - lateral elbow pain with resisted wrist extension performed with the elbow extended, forearm pronated and wrist flexed
    - tenderness on palpation at or close to the lateral epicondyle
    - lateral elbow pain with resisted middle finger extension
    - lateral elbow pain with passive overpressure with the forearm held in a position of full elbow extension, full wrist flexion, pronation and ulnar deviation.
3. no history of fractures, elbow surgery or osteoarthritis
4. no active shoulder or wrist tendinitis
5. no systemic diseases or disorders such as rheumatoid arthritis or fibromyalgia

6. ability to understand spoken and written English

7. ability to complete questionnaires independently

8. access to a telephone

9. provision of written consent (Appendix H)

All eligible subjects were provided with an information letter describing the study (Appendix I). If the subjects indicated on the information letter they wanted to volunteer for the study their name and telephone number were forwarded to the principal investigator who contacted the subjects and arranged to meet back at the clinic to collect data within the week.

During the eight-month period between June 1997 and January 1998 a total of thirty-two subjects (18 women and 14 men) volunteered. Eight people refused to participate, either they did not want to be involved in a research study (4), had other time restrictions (3), or had inadequate funds to pay for a course of treatment (1).

Outcome Measures

This study used six questionnaires, the Tennis Elbow Function Scale (TEFS), Pain Free Function Index (PFF), Visual Analogue Scales for Pain (pain VAS) and Function (function VAS), and change rating scales for therapist (therapist CRS) and subject (subject CRS). Two performance-based outcome measures, pain free grip strength (PFG) and pain free grip strength on the unaffected arm (MGU), were also used.

Both the PFF and TEFS questionnaires have previously been described (Appendices A and B). The total score was the criterion measure for each scale.

Function was also measured using the function VAS (Appendix J). The score was the distance, in centimetres, from the left end of the ten-centimetre horizontal line. Pain was measured using the pain VAS (Appendix J). The scoring for the pain VAS was the same as for the function VAS.

In order to establish a standard criterion for identifying those subjects who experienced a clinically significant change in symptoms the Change Rating Scale, an ordinal scale with the categories of worse, same, and improved, was developed for the study. There were two versions of the CRS, one for the subject, (subject CRS-Appendix K) and the other for the therapist (therapist CRS-Appendix L). When both the subject and therapist agreed on a rating of improvement, the subject was considered improved and data from the subject were included in the appropriate responsiveness analysis.

The CRS scale was modeled after similar scales used to determine the responsiveness of new measures.[46, 58] Previously, a three-point scale with similar wording was used to evaluate the responsiveness of the Shoulder Pain and Disability Index.[46] A five-point rating of change scale was used in a study investigating the efficacy of ultrasound treatment for LE, and correlated with PFG and PFF ($r = 0.51$ and $r = 0.57$).[54] In that study subjects were required to rate whether their symptoms were much worse, slightly worse, unchanged, slightly better or much better. Because there was a small sample size, data were later reduced to four categories to facilitate the analyses of subjects' change scores.[54] In the present study, the purpose was to identify subjects who demonstrated clinical improvement; and it was thought that three categories would be adequate to identify improved subjects.

Grip strength was used in this study because it is a valid, reliable and responsive outcome measure for LE and because it measures physical impairment.[9] The test position, procedure and timing of measures used in this study were similar to those used in another LE study which examined the reliability of pain free and maximum grip strength tests.[56] The same Smedley[1] type Jamar™ dynamometer was used for all subjects and was calibrated against a known set of weights prior to use in the study. The subjects stood holding the grip dynamometer with the elbow extended and the shoulder and forearm in neutral rotation. To test pain free grip the subjects squeezed the dynamometer with the involved limb until they first experienced pain. The force reading (kg) on the dynamometer at this point was recorded. PFG was calculated by averaging the values of three consecutive repetitions separated by a 20-second rest. The same procedure was used to test grip strength on the unaffected arm except that the force recorded represented maximum grip strength (MGU).

Data Collection

The first meeting between the subject and principal investigator occurred within one week of the subject's volunteering for the study. At this meeting, Test 1, the principal investigator confirmed that the subject met the inclusion criteria. The purpose of the study was discussed and any concerns the subject had regarding the study were addressed. Subjects were also requested to refrain from changing their medication regimen and using a forearm strap until the second test had been completed. Then the consent form was signed.

---

[1] Therapeutic Equipment Corporation, Clifton NJ.

The descriptive data were obtained for the following variables: gender, age, symptom duration, dominance, and affected limb. Subjects were shown the pain VAS, the function VAS, the PFF, and the TEFS. The purpose of each questionnaire and the procedure to complete each questionnaire were discussed. Subjects were left alone in a quiet area to complete the questionnaires, but the principal investigator was available to assist subjects with this task. PFG and MGU were then tested.

Within four working days of Test 1, subjects were asked to complete the PFF and TEFS questionnaires again. If Test 2 was scheduled at the same time as a physical therapy visit, administration of the indices took place prior to the subjects' receiving therapy.

The third visit was scheduled at the clinic six weeks following Test 2. The principal investigator showed subjects the (subject) CRS and asked them to complete it. Then all tests were repeated and the number of treatments was recorded. Following this, the therapist responsible for treating the subject reexamined the subject and completed the (therapist) CRS. This therapist was blinded to the results of the (subject) CRS.

Analysis

Inferential statistics were calculated at a significance level of 0.05. Descriptive statistics for age, gender, duration of symptoms, limb dominance, affected limb and total number of treatments were calculated. Means and standard deviations were also calculated for each self-report outcome measure, grip strength measures and for difference scores of measures taken on different occasions.

Test retest reliability of the TEFS was determined by calculating the ICC (1,1)[61] and confidence intervals.[62] The standard error of measurement (SEM) was also calculated [11] (Appendix M).

Construct validity of the TEFS was determined by calculating the Pearson product-moment correlation coefficients (one-tailed) for the Test 1 scores of the TEFS and the other outcome measures. The Pearson correlation coefficient was chosen over Spearman rank correlation coefficient because all measures except the PFF and TEFS have previously been treated as continuous variables.

Responsiveness was determined using four statistical analyses (Appendix N). First, a repeated measures ANOVA was used to identify differences between the TEFS scores of Test 1, Test 2 and Test 3. If a significant difference existed, Tukey's Honestly Significant Difference was used to determine where differences occurred. Second, the Relative Efficiency[50] [paired t TEFS/paired t PFF]^2 was calculated for subjects who had improved. Third, the Responsiveness Index[47, 49] for the PFF and TEFS was calculated and compared for the entire group and for those rated as improved. Finally, using data from improved subjects, the Pearson correlation coefficients were correlated for the change scores of the TEFS and other measures.

Ethical Considerations

This study received ethical approval from the Health Research Ethics Administration Board Committee B (Appendix O) and administrative approval from Caritas (Grey Nuns Hospital) (Appendix P). The TEFS was copyrighted and one of the developers, Paul Stratford, provided written consent for the TEFS to be used for

research.[57] All participating clinics provided written permission for the researcher to recruit patients from their sites.

Eligible subjects provided written permission to be contacted by the principal investigator. It was explained that this study presented no risk to subjects because there were no reported injuries resulting from grip strength testing. Furthermore, the possibility of harming a subject was low because the subject stopped the grip test when he or she felt pain. Treatments were not administered by the principal investigator, therefore the liability and risks associated with providing treatments remained in the domain of the referring clinic. It was acknowledged that the extra test visits to the physical therapy clinic might have inconvenienced the patient.

Written consent was obtained from all subjects prior to their participation in the study. Subjects were informed they could withdraw from the study without any consequences to receiving more physical therapy treatment. However, the subjects were informed the principal investigator would contact those who dropped out to determine the reason for stopping participation. All participants were given the opportunity to request a copy of any publications arising from this study.

Chapter 4

Results

Subject Characteristics

Thirty-one of the 32 subjects enrolled in the study were included in the final analyses. One person was dropped from the study after she developed an infection in her arm. Ages of the 31 subjects ranged from 34 to 63 years ($\bar{X}$= 44.71 $\pm$ 7.92). Duration of symptoms ranged from 1 week to 15 years. Eight subjects had experienced continuous symptoms for one year or greater. The average duration of symptoms for the other 24 subjects with symptoms less than one year was 10.75 $\pm$ 3.28 weeks. Three subjects considered this episode a recurrence following a period of being symptom free. Ninety percent (29) of the subjects were right limb dominant. The dominant limb was affected in 71% (22 right and 1 left) of the subjects.

Prior to the initial assessment subjects received an average of one treatment. Between the initial and the second test sessions, 13 subjects received an additional treatment. One subject received two additional treatments.

The total number of treatments administered during the six to eight week period varied from 2 to 18. An average of 7.61$\pm$3.75 treatments was administered to the 31 subjects remaining in the study at the final test session.

At the final test session, 25 subjects were considered improved because they were rated improved by both subject and clinician on the Change Rating Scales. Five of the

improved subjects considered their symptoms resolved. The remaining 6 subjects received a rating of unchanged and were considered not improved. Three of these subjects had discontinued their treatment and were awaiting assessment from orthopaedic surgeons. Two subjects had used their eligible physical therapy coverage (9 or 10 visits) but felt that the condition had not improved. One subject discontinued treatment after two visits because he could not afford the treatments.

Analysis of the TEFS

Tables 1 and 2 contain the descriptive statistics for scores of all measures and the difference scores for the measures. The ICC (1,1) for the TEFS test retest reliability was 0.92 (95% confidence interval = 0.83-0.96). The SEM for the TEFS was $\pm$ 2.35.

The Pearson correlation coefficients in Table 3 indicate the TEFS is moderately correlated with the PFF, the function VAS, the pain VAS and PFG. MGU was not related to the TEFS.

Table 1: Descriptive Statistics for Outcomes Measures taken on the Three Test Sessions for all Subjects (n=31)

	Test 1		Test 2		Test 3	
	Mean	SD	Mean	SD	Mean	SD
TEFS *	23.06	7.90	22.48	8.25	11.55	9.85
PFF **	6.42	1.20	6.48	1.43	4.26	2.67
Function VAS (cm)	4.77	2.37			2.33	2.13
Pain VAS (cm)	5.92	2.31			2.84	2.41
PFG (kg)	13.89	9.58			22.63	11.17
MGU (kg)	31.66	9.19			32.56	10.09

* TEFS total possible score = 40

** PFF total possible score = 8

Table 2: Descriptive Statistics for the Difference Scores of Outcome Measures for all Subjects (n=31)

	Test 1-3 difference		Test 1-2 difference		Test 2-3 difference	
	Mean	SD	Mean	SD	Mean	SD
TEFS	11.52	6.87	0.58	3.32	10.94	8.07
PFF	2.16	2.54	-0.06	0.99	2.23	2.26
Function VAS (cm)	2.45	1.85				
Pain VAS (cm)	3.07	2.09				
PFG (kg)	-8.75	7.66				
MGU (kg)	-0.89	4.18				

a negative sign (-) denotes either a lower PFF score or greater grip strength at Test 1 vs. Test 2 or Test 3

Table 3: Pearson Correlation Coefficients and p-values () for Measures at Test 1 (n=31)

		TEFS	PFF	Function VAS	Pain VAS	PFG
PFF	r	**.47**				
	p	(.00)				
Function VAS	r	**.74**	**.39**			
	p	(.00)	(.02)			
Pain VAS	r	**.51**	.30	**.64**		
	p	(.00)	(.05)	(.00)		
PFG	r	**-.47**	**-.45**	-.30	**-.33**	
	p	(.00)	(.01)	(.05)	(.04)	
MGU	r	-.22	**-.31**	-.17	.05	**.53**
	p	(.11)	(.04)	(.36)	(.39)	(.00)

Significant correlations ($p < 0.05$ one–tailed) are highlighted in bold

Table 4: Comparison of TEFS Scores at Test 1,2, and 3

Test Pairs	Difference in Means
T1-T2	0.58
T1-T3	**11.52**
T2-T3	**10.94**

2.946 is the value required for the difference in means to be statistically significant ($F = 63.42$, $p = .00$)

The repeated measures ANOVA indicated that there was a significant difference between the TEFS scores across time. Post hoc analysis indicated that Test 3 was significantly different from Test 1 and Test 2 (Table 4).

The Relative Efficiency for improved subjects was 4.25 $[t\ TEFS/t\ PFF]^2$ (Appendix N). A Relative Efficiency greater than one indicates that the TEFS scores changed more than the PFF scores.

The TEFS Responsiveness Index for the entire group and for the subset of subjects rated improved was greater than the PFF Responsiveness Index for both groups (Table 5).

Table 5: Responsiveness Index for All Subjects and Improved Subjects

	All Subjects n=31	Improved Subjects n=25
TEFS	3.30	4.22
PFF	2.25	2.61

The Pearson correlation coefficients for changes in scores indicate there is a fair relationship between changes in the TEFS and changes in the PFF, function VAS and pain VAS. There was no relationship between changes in TEFS and grip strength (Table 6).

Table 6: Pearson Correlation Coefficients and p-values () for Change Scores (n=25)

		TEFS	PFF	Function VAS	Pain VAS	PFG
PFF	r	**.37**				
	p	(.03)				
Function VAS	r	**.40**	.14			
	p	(.02)	(.24)			
Pain VAS	r	**.39**	.00	**.37**		
	p	(.03)	(.41)	(.03)		
PFG	R	-.11	**-.50**	-.05	-.07	
	P	(.30)	(.01)	(.41)	(.38)	
MGU	R	.03	-.25	.20	**.49**	-.08
	P	(.44)	(.12)	(.17)	(.01)	(.35)

Significant correlations (p. < 0.05 one-tailed) are highlighted in bold

The distribution of Test 1 TEFS scores appears to be normal (Figure 1). The distribution of Test 1 PFF scores appears to be skewed to the right (Figure 2).

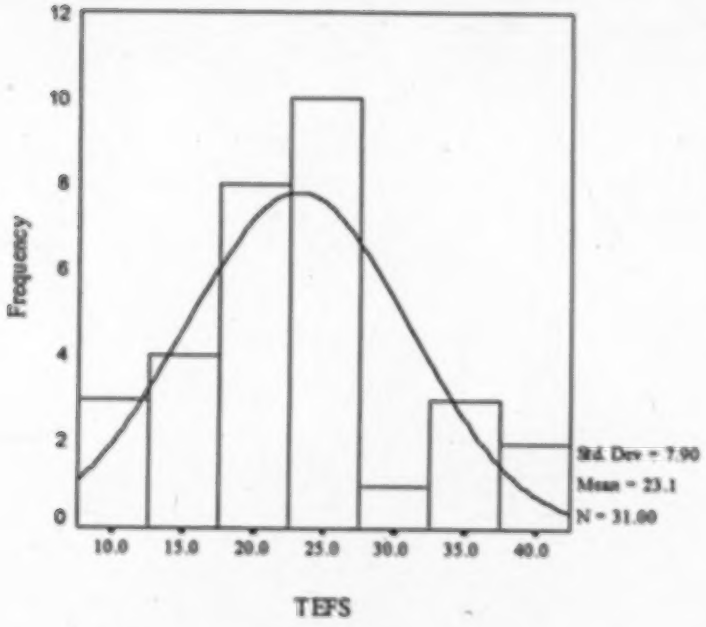

Figure 1: Distribution of Test 1 TEFS scores

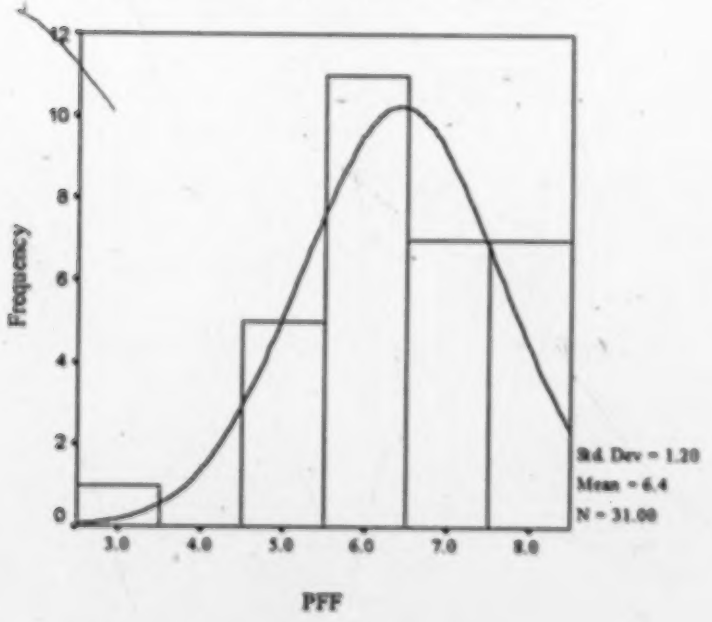

Figure 2: Distribution of Test 1 PFF Scores

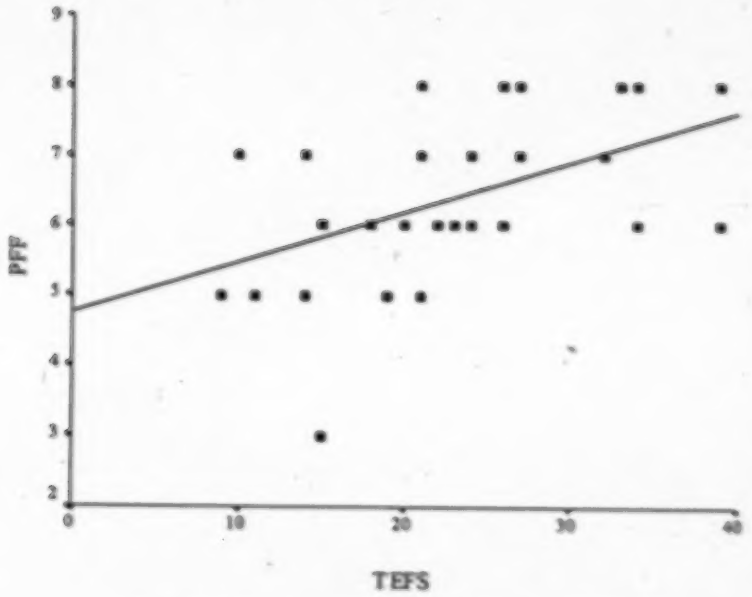

Figure 3: Correlation of the Test 1 TEFS and PFF

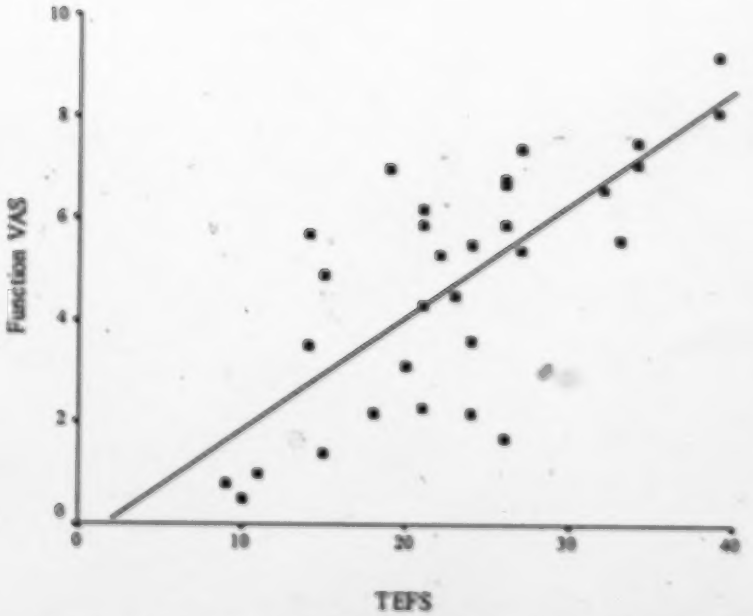

Figure 4: Correlation of the Test 1 TEFS and Test 1 function VAS

Figure 3 illustrates the correlation between Test 1 TEFS and PFF measures.

Figure 4 illustrates the correlation between Test 1 and TEFS and Function VAS.

Chapter 5

Discussion

The purpose of this study was to examine the psychometric properties of the TEFS, a self-report measure of pain-related function in subjects with LE. The main objectives were to examine the test retest reliability, the construct validity and the responsiveness of the TEFS.

This study found that the TEFS is a reliable, valid and responsive measure and should be considered suitable for evaluating patients with LE. A discussion concerning the generalizability of results, the reliability, construct validity, responsiveness of the TEFS, and general observations follows.

Generalizability of Results

Subjects in this study were similar to those in other studies of LE with respect to average age, age range and duration of symptoms. Like other studies, at least 20% of the sample population experienced symptoms for more than one year. This indicates that for a considerable proportion of this population, the condition may not be self-limiting.[18, 24, 31] The dominant limb was affected in 71% of the subjects. This proportion is 10% lower than that reported by Stratford et al.,[17] but similar to that reported in other clinical trials.[31, 38] This difference may be attributed to the type of clinic used in the studies. A single sports injury clinic was used by Stratford et al.[9], whereas, multiple hospital-based and private physical therapy clinics were used in this and other studies. The sports injury

clinic may have attracted a greater number of athletes whose increased activity level may have made them susceptible to developing LE in their dominant limb. Overall the subjects in this study appear similar to samples in other studies [18, 31] and the results should generalize to patients with LE who seek physical therapy care and do not have the comorbid conditions noted as exclusion criteria.

The average number of treatments subjects received (7.61 ± 3.75) is probably less than what is usually provided for this condition because the final test occurred prior to the completion of therapy for 52% (16) of the subjects. Other subjects (3 improved, 2 not improved) limited their physical therapy visits to the maximum allowed by their health insurance plan or the regional rehabilitation program (9 or 17). In five subjects the condition resolved by the final testing date (number of visits ranged from 2-9).

Reliability

The test retest results of the ICC (0.92), confidence interval (95% CI 0.83-0.96) and SEM (± 2.35) suggest that the TEFS is a reliable measure. The results of the ANOVA and post hoc analysis of difference in mean scores of the TEFS also indicate that it is a reliable measure.

As hypothesized the reliability coefficient (ICC) for the TEFS in this study was high. A high ICC indicates that the variance within subjects (difference between Test 1 and 2) was small compared to the variance between subjects. With reliability coefficients, there are varying opinions as to what is considered an acceptable level for self-report measures.[11, 63] Portney et al.[11] suggest that measures with reliability coefficients above 0.75 have good reliability and those with coefficients below 0.75 have

poor to moderate reliability. A measure with a reliability coefficient of 0.75 may be acceptable for studies using large sample sizes.[64] However, when measures are used to monitor an individual's response to treatment, the reliability coefficient should exceed 0.90.[11] The 95 % confidence interval, a calculation of measurement error for the ICC, indicated that the true value of the TEFS reliability coefficient falls within a range considered high. Thus the analysis of reliability using the ICC and its CI indicates the TEFS has good reliability and should be suitable for clinical trials investigating LE and for assessing within patient changes.

The standard error of measure (SEM) represents the population of all possible measurement errors that could occur for a variable.[11] The SEM is a statistic that is useful for clinicians because it is expressed in the same units as measured by the TEFS. The results indicated that on repeated administration of the TEFS, there was a 68% chance that the true score of the TEFS was $\pm$ 2.35 of the measured value. When the SEM is doubled (i.e. within 2 standard deviations of the mean score) there is a 96 % chance the true score lies within $\pm$ 4.7 of the measured value. Thus if the TEFS scores change $\pm$ 4.7 the change likely reflects a real change in pain-related elbow function.

The repeated measures ANOVA and post hoc analysis of TEFS mean scores at Test 1, 2 and 3 (Table 4) indicated there were no significant differences in mean scores between Test 1 and Test 2. The difference in mean Test 1 and Test 2 TEFS scores (0.58) was considerably lower than 2.946, the value at which difference in means would be significant. This indicates that there was no significant change in subjects' scores between Test 1 and 2, further supporting the reliability of the TEFS.

Several factors may have contributed to the stability of test conditions and the little variation between subjects' scores on repeat TEFS administration. Every attempt was made to standardize the testing conditions and to minimize treatment between Test 1 and Test 2. One therapist, the principal investigator, administered the TEFS to each subject on repeated occasions. Subjects were tested in a quiet area in the same clinic and were requested to refrain from changing their medication regimen or using a forearm strap until after the second test session. Furthermore, it appears that the condition of LE did not change significantly in the first week of treatment. Although fourteen subjects received treatment between Test 1 and Test 2, the ICC was high. This finding is consistent with another study in which the condition was considered stable in the first week of treatment because there were high reliability coefficients for pain, grip strength and function measures.[9]

Construct Validity

The results support that the TEFS is a valid measure for LE because the TEFS demonstrated fair to good relationships with other valid clinical measures for LE (Table 3). The construct validity of the TEFS as a measure of pain-related elbow function was also supported because the TEFS was related to measures of similar constructs, pain intensity and function (convergent construct validity), and because the TEFS was not related to a measure of an unrelated construct, maximum grip strength (divergent construct validity).[65]

The correlation between the TEFS and the pain VAS was slightly less than predicted (0.49 vs. > 0.50), however, this difference is negligible and the hypothesis that

the TEFS and pain VAS are moderately positively correlated is accepted. One would expect a moderate relationship between pain intensity and discomfort-related elbow function because pain with gripping activities is the prime symptom of LE.[10] The validity of the TEFS as a measure of pain-related function is supported because of this correlation with a pain intensity measure.

A moderate correlation of 0.49 between the TEFS and pain VAS indicates that only 24% of the variability in pain intensity as measured by the VAS accounts for the variability in discomfort-related function as measured by the TEFS. A number of other factors such as subjective perception of the effort required to complete the tasks listed, type of task, pain beliefs, fear-avoidance behavior, grip strength, and muscle endurance may also contribute to function.[45, 66, 67]

The TEFS was strongly related to the function VAS. This relationship suggests that the TEFS measures elbow function. Intuitively, it was hypothesized that the TEFS would demonstrate a stronger correlation with the PFF ($r = 0.47$), a multi-item measure of the construct of pain-free elbow function, than with the function VAS ($r = 0.74$), a single-item measure of overall elbow function. The hypothesis was not supported. This was unexpected because some activities listed on the TEFS are the same on the PFF.

The low correlation between pain free function and pain-related function may be the result of a ceiling effect associated with the PFF. The PFF scale requires patients to indicate if pain is present or absent. Therefore, PFF scores will be high, for patients whose pain intensity is minimal or severe (Figure 2). The TEFS has an ordinal scale with a greater possible range of scores that subjects can choose from to reflect variations in

their pain intensity. The TEFS scores will be low for subjects with minimal pain symptoms and high for those with severe pain symptoms. The correlation between the TEFS and the PFF was low because mild funtional difficulties due to pain as recorded by the TEFS would be recorded as more severe difficulty by the PFF (Figure 3).

To further examine if a ceiling effect was associated with the use of the PFF the standard deviation and coefficient of variation[11] was calculated (CV= SD/X x 100). Examination of the standard deviation of the Test 1 function measures (Table 1) suggests that the PFF scores do not vary greatly. The lower coefficient of variation for the PFF (18%) compared to the function VAS (48%) and TEFS (35%) indicates that in this sample the PFF scores were less variable. The lower variability demonstrates that the PFF scores were clustered.

The function VAS has a continuous scaling format, whereas, the TEFS has an ordinal scaling format. The coefficients of variation for these measures (48% and 35% respectively) demonstrated that there was a similar degree of variability in each measure. Both outcomes provided subjects with a range of scores from which they could choose to represent varying levels of pain-related function. Scores for the subjects with minimal difficulty with pain-related function would be lower than scores for the subjects who had more difficulty with pain-related function. Thus a stronger correlation resulted between the TEFS and the function VAS than between the TEFS and the PFF (Figure 4). The data are in concordance with Streiner and Norman's [64] observations that when continuous variables (function) are represented by a dichotomous scale (Painfree function scale),

there is a loss of efficiency in the instrument and the correlation with other measures is generally lower.

The strength of the correlations with both the function VAS and the PFF indicates that the TEFS does measure aspects of elbow function in subjects with LE. The moderate correlations suggest that all three scales measure unique aspects of elbow function in LE patients.

There was a moderate negative relationship between the TEFS and pain-free grip strength on the affected limb. One would have expected pain-related function and pain free grip strength to be related because the TEFS test items were generated by patients who had LE, and difficulty gripping is a key complaint of patients with LE.[57] The negative value indicates that higher scores on the TEFS (more difficulty with function) were associated with weaker pain-free grip strength. Other painful upper extremity musculoskeletal conditions also demonstrate similar relationships between increased pain intensity, decreased function and reduced grip strength.[45, 68]

As predicted, the TEFS and MGU were not related. Grip strength measures on the unaffected arm were less likely to be affected by pain and more likely attributed to gender, age, activity levels and anthropometric properties such as height, weight and fat-free muscle mass.[69]

Responsiveness

The analysis indicated that the TEFS is responsive and more efficient than the PFF at measuring change. The Pearson correlations between the TEFS change scores and

other measures support the validity of the TEFS as being sensitive to changes that occurred in LE. However, the correlations between measures were not as strong as hypothesized.

The first analysis was to determine if TEFS scores changed over time. The greatest change in TEFS scores occurred between Test 1 and 3 when subjects received the most treatments and the most time passed between measures. It was assumed that the treatments and the passage of time would act as agents for decreasing pain-related elbow function. As predicted, these changes in function were reflected in the TEFS scores.

The TEFS was a more responsive measure than the PFF when compared using Relative Efficiency. The RE score of greater than one indicates that the t-score of the numerator, the TEFS, was larger than the t-score of the denominator, the PFF. On the PFF, one must experience a complete resolution of discomfort while performing a specified activity before the score of a PFF test item will change. In contrast, smaller changes in pain-related function can be documented on the TEFS because there is a five-point ordinal scale.

The Responsiveness Index (RI) [49] of the TEFS was greater than that of the PFF for the entire group and for improved subjects. The higher RI for the TEFS indicates it recorded a greater change in scores than the PFF. This result can be explained by the scaling format of the TEFS. There was a greater potential for changes in the condition to be reflected in scores on the TEFS, because the total possible score for the TEFS was 40, whereas the score for the PFF was only 8.

The final analysis found the changes in TEFS were related to changes in the pain VAS, function VAS and PFF. The relationships were lower than the predicted correlation of ($r \geq 0.5$). Despite this, the validity of the TEFS to measure change is supported because it related to other measures that already have been proven to be responsive to changes in LE.

Changes in pain free grip strength on the affected arm were not related to changes in the TEFS. The self-report scales are subjective measures of pain intensity, function or pain-related function, whereas pain-free grip strength is an objective performance-based physiological measure. It was assumed resolution of the condition as measured by the TEFS would correspondingly be reflected in pain free grip strength. However other factors may have affected subjects' estimation that their pain-related function had improved even when pain free grip strength had not.

Treatment may have resulted in decreased pain intensity or subjects may have also been advised to, or intuitively, limited the activities that caused their pain. The inactivity related to treatment may have led to deconditioning. Furthermore, except for the first two items, most of the tasks on the TEFS appear to require a limited amount of grip strength. A small decrease in pain intensity may have resulted in the perception that pain-related function for the easier tasks improved significantly without a corresponding increase in pain free grip strength. Several subjects noted on the final test that their pain had decreased and they thought they were improved. However, when they performed more strenuous tasks such as sweeping or shoveling snow, they experienced an exacerbation of symptoms. Two subjects performed strenuous tasks the day before the final test session. They thought pain with gripping had increased since the previous day.

The lack of a relationship between changes in the TEFS and pain free grip strength indicates that both measures would be useful as outcomes to measure changes in the condition because they reflect different aspects of recovery.

General Observations

It was anticipated there would be some difficulty with subject recruitment, on account of how physical therapy was funded in the Capital Health Region. The Community Rehabilitation Program requires that therapists use a Determination of Needs form to rate if individuals qualify for public funding. The assessment tool asks the physical therapists to rate the functional impact of the musculoskeletal problem, the predicted outcome if treatment is administered and the risk of the condition becoming worse if treatment is delayed. A minimum score of seven out of a total possible score of fifteen is required for a patient to qualify for funding. If clients do not qualify for funding they are required to pay for their treatment. The use of the Determination of Needs form may have caused subjects with minimal symptoms of LE to be considered ineligible for publicly funded care and subsequently they were unable to pay for their therapy. At least one subject agreed to be in the study but was unable to volunteer once she found she did not qualify for public funding. It is not known exactly how many of the other subjects were determined ineligible for CRP funding and required to pay for treatment using their extended health benefit plan or their own funds.

Access to adequate funds for treatment may have affected the improvement patients experienced because at least two subjects stopped treatment once they used their allotment of publicly funded care. The average number of treatments subjects required to

the resolve this condition is not known because final testing occurred before most subjects' final physical therapy visit.

The researcher chose to use multiple data collection sites because few clinics treated more than five LE patients per month and it was anticipated there would be difficulty recruiting an appropriate sample size within a reasonable time frame. Initially five clinics were asked to serve as referral sites and it was hoped that data collection would be completed within three months. However data collection at three months was slower than anticipated, so five other clinics were asked to act as referral sites. Despite the addition of the extra five clinics it took eight months to recruit an adequate sample size.

Stratford et. al.[54] notes that most clinical trials for LE lack the power to detect a treatment effect because there is an insufficient sample size. One explanation for this may be that the incidence of isolated cases of LE is lower than reported. When several therapists were questioned if all patients with LE were asked to participate in the study the therapists noted that they had patients with LE on their caseload, however those patients were excluded because they also had wrist or shoulder problems. In this study the number of patients with both LE and associate wrist and shoulder problems was not determined.

Another explanation may be related to the nature of physical therapy service delivery in the region were the studies are conducted. In the Capital Health Region, if patients know they must undergo an assessment to determine their eligibility for physical therapy, they may delay seeking physical therapy treatment until the problem is severe or multiple structures are affected. If the access to and availability of physical therapy

funding is limited for people with LE, then subject recruitment for studies may be difficult. A central site for patient referral and unrestricted funding for all cases regardless of the severity of the condition may facilitate recruitment of greater numbers of subjects, thus improving the power of clinical trials to detect treatment effects.

Both the PFF and TEFS ask the subject to rate how much discomfort he or she experience when performing activities. Several subjects noted that as their symptoms improved they no longer experienced elbow pain or discomfort, but they experienced a sensation of elbow stiffness. Subjects did not think stiffness was a concept that was addressed by any questionnaire in this study. In the published literature perception of stiffness has not been documented or measured. Decreased wrist range of motion has been noted in those with LE and symptoms have been decreased in response to a forearm-stretching program for LE.[6, 19] There may be a need to measure perception of stiffness and wrist range of motion in LE subjects.

In this study, using the Change Rating Scale, improvement was linked to the therapists' and subjects' perception that the condition had improved. The use of the change rating scale facilitated the responsiveness analysis. However, the reliability and validity of the Change Rating Scale had not been established.

There may also be bias inherent in using the therapist Change Rating Scale because the therapists based their judgement on patient reports of pain intensity and the results of a clinical reexamination of the elbow. The therapists' judgement may have been confounded by the inability to recall the results of the initial clinical examination and by a reliance on subjects' pain and function assessments. Clinician judgement of disease severity in arthritis patients was linked more closely with the patients' assessment

of pain than with the objective assessment of disease activity.[70, 71] This study could have been improved by administering the Change Rating Scale on the three test occasions, and examining its reliability and validity.

This study used subject and therapist assessment of change to identify improved subjects. Another study of LE linked improvement in the clinical condition with the ability to progress to pain-free wrist extensor strengthening exercises.[9] Subsequent clinical trials for LE may examine whether performance-based measures or self-report measures are better predictors of a positive response to treatment.

Suggestions for Future Research

The TEFS is a valid and responsive measure of pain-related function for LE. Because it has been validated, the TEFS represents an improvement compared to the other measures of function reported in the published LE clinical trials. If the TEFS was used routinely in subsequent clinical trials, responsiveness would be enhanced and comparison of the results from those trials would be facilitated.

The results indicated the TEFS measures a different construct from those measured by the other self-report questionnaires and pain free grip. Several subjects noted that as they improved and their pain and discomfort resolved, they experienced stiffness of the elbow musculature. They did not think the TEFS or any of the other outcome measure adequately tracked this construct. Future studies may choose to also measure elbow stiffness as an outcome for LE.

The validity and reliability of the change rating scale was not established in this study. A rating of change scale with more cut points would be useful in subsequent LE

clinical trials because a more expansive change rating scale could be used to identify if the effects of treatment differs for patients.

Conclusion

This study analyzed the measurement properties of the Tennis Elbow Function Scale (TEFS). The results indicate that the TEFS is a reliable, valid and responsive outcome measure suitable for evaluating changes in pain-related function in subjects with LE. The reliability of the TEFS is considered good because there was a high ICC for repeated measures administered during the first week of treatment.

The construct validity of the TEFS is supported because it correlated with related measures of pain intensity, elbow function, pain free elbow function and pain free grip strength in the affected arm. The TEFS was sensitive to changes in subjects who received physical therapy and who were assessed as improved at the end of a seven-week period. Furthermore, the TEFS was found to be more responsive when compared to the Painfree Function Index (PFF). A ceiling effect was noted with the PFF. Scores were high even for subjects with minimal symptoms so the PFF may not reflect small decreases in subjects' pain level and related improvements in function. Change scores of the TEFS were fairly to moderately correlated with the change scores of other questionnaires for pain intensity, pain free function and overall function. There was no correlation with changes in pain free grip strength. Researchers should consider using a combination of the self-report questionnaires and pain free grip strength when evaluating recovery in subjects with LE because the outcomes measure different aspects of the condition. The validation of the TEFS represents a substantial advancement for

clinicians and researchers of LE because the TEFS appears to be more responsive than the PFF. Future research could be improved by standardized use of the TEFS, which would facilitate comparisons between clinical trials.

References

1. Allander E. Prevalence, incidence and remission rates of some common rheumatic diseases and syndromes. Scandinavian Journal of Rheumatology 1974; 3:145-153.
2. Priest J, Braden J, Gerberich J. The elbow in tennis, Part 1: an analysis of players with and without pain. Physician and Sports Medicine 1980; 8:89-91.
3. Labelle H, Guibert R, Joncas J, Newman N, Fallaha M, Rivard C. Lack of scientific evidence for the treatment of lateral epicondylitis of the elbow. Journal of Bone and Joint Surgery[Br] 1992; 74-B:646-651.
4. Hudak PL, Cole DC, Haines T. Understanding prognosis to improve rehabilitation: The example of lateral elbow pain. Archives of Physical Medicine and Rehabilitation 1996; 77:586-93.
5. Solveborn SA, Buch F, Mallmin H, Adalberth G. Cortisone injection with anesthetic additives for radial epicondylagia (Tennis Elbow). Clinical Orthopaedics and Related Research 1995; 316:99-105.
6. Solveborn SA. Radial epicondylalgia ('tennis elbow'): treatment with stretching or forearm band. A prospective study with long-term follow-up including range-of-motion measurements. Scandinavian Journal of Medicine and Science in Sports 1997; 7:229-237.
7. Labelle H, Guibert R. Efficacy of diclofenac in lateral epicondylitis of the elbow also treated with immobilization. Archives of Family Medicine 1997; 6:257-62.
8. Pienimaki TK, P. Kemila,T. Koivukangas, P. Vanharanta, H. Long-term follow up of conservatively treated chronic tennis elbow patients. A prospective and

retrospective analysis. Scandinavian Journal of Rehabilitation Medicine 1998; 30:159-166.

9. Stratford P, Levy DR, Gauldie S, Levy K, Miseferi D. Extensor carpi radialis tendonitis: a validation of selected outcome measures. Physiotherapy Canada 1987; 39:250-255.

10. Reid DC. Sports Injury Assessment and Rehabilitation. New York: Churchill Livingstone, 1992:1013 - 1022.

11. Portney L, Watkins M. Foundations of Clincal Research: Applications to Practice. East Norfolk: Appelton and Lange, 1993.

12. Guyatt G, Feeny DH, Patrick DL. Measuring health-related quality of life. Annals of Internal Medicine 1993; 118:622-629.

13. Nirschl R. Elbow Tendinosis/Tennis Elbow. Clinics in Sports Medicine 1992; 11:851-870.

14. Binder AI, Hazleman BL. Lateral humeral epicondylitis - a study of natural history and the effect of conservative therapy. British Journal of Rheumatology 1983; 22:73-76.

15. Gruchow HW, B.S. P. An epidemiological study of tennis elbow. American Journal of Sports Medicine 1979; 7:234 -248.

16. Gellman H. Tennis Elbow (lateral epicondylitis). Orthopedic Clinics of North America 1992; 23:75-82.

17. Stratford P, Levy DR, Gauldie S, Miserferi D, Levy K. The evaluation of phonophoresis and friction massage as treatments for extensor carpi radialis tendonitis: a randomized controlled trial. Physiotherapy Canada 1989; 41:93-99.

18. Holdsworth LK, Anderson DM. Effectiveness of ultrasound used with a hydrocortisone coupling medium or epicondylitis clasp to treat lateral epicondylitis: pilot study. Physiotherapy 1993; 79:19-25.

19. Solveborn S, Olerud C. Radial epicondylgia measurement of range of motion of the wrist and elbow. Journal of Orthopaedic and Sports Physical Therapy 1996; 23:251-257.

20. Runge R. Zur Genese und Behandlung des Schreibkrampfes. Beitrage 1873; 10:245-248.

21. Viikari - Juntura E, Kurppa K, Kuosma E, et al. Prevalence of epicondylitis and elbow pain in the meat processing industry. Scandinavian Journal of Work and Environmental Health 1991; 17:38-45.

22. Luopajarvi T, Kuorinka I, Virolainen M, Holmberg M. Prevalence of tenosynovitis and other injuries of the upper extremities in repetitive work. Scandinavian Journal of Work Environment and Health 1979; 5:48-55.

23. Pheasant S. Ergonomics, Work and Health. Gaithersburg, Maryland: Aspen Publishing Inc., 1991:88-97.

24. Cyriax J. Textbook of Orthopaedic Medicine. Vol. one. London: W.B. Saunders, 1982.

25. Boyd HB, Mcleod AC. Tennis Elbow. Journal of Bone and Joint Surgery (AM) 1973; 55A:1183-1187.

26. Clarke A, Woodland J. Comparison of two steriod preparations used to treat tennis elbow, using the hypospray. Rheumatology and Rehabilitation 1975; 14:47-49.

27. Gerberich SG, Priest JD. Treatment for lateral epicondylitis: variables related to recovery. British Journal of Sports Medicine 1985; 19:224-227.

28. Dijs H, Mortier G, Driessens M, De Ridder A, Willems J, De Vroey T. A retrospective study of the conservative treatment of tennis elbow. Medica Physica 1990; 13:73-77.

29. Lundeberg T, Abrahamsson P, Haker E. A comparative study of continuous ultrasound, placebo ultrasound and rest in epicondylagia. Scandinavian Journal of Rehabilitation Medicine 1988; 20:99-101.

30. Devereaux MD, Hazleman BL, Thomas P. Chronic lateral epicondylitis-a double-blind controlled assessment of pulsed electromagnetic field therapy. Clinical Experiments in Rheumatology 1985; 3:333-336.

31. Pienimaki TT, Tarvainen TK, Siira PT, Vanharanta H. Progressive strengthening and stretching exercises and ultrasound for chronic lateral epicondylitis. Physiotherapy 1996; 82:25-33.

32. Haker E, Lundeberg T. Acupuncture treatment in epicondylalgia: a comparative study of two acupuncture techniques. The Clinical Journal of Pain 1990; 6:221 - 226.

33. Burton A. Grip strength and forearm straps in tennis elbow. British Journal of Sports Medicine 1985; 19:37-38.

34. Haker E, Lundeberg T. Pulsed ultrasound treatment in lateral epicondylalgia. Scandinavian Journal of Rehabilitation Medicine 1991; 23:115-118.

35. Brattberg G. Acupuncture therapy for tennis elbow. Pain 1983; 16:285-288.

36. Chalmers TC, Smith HJ, Blackburn B. A method for assessing the quality of a randomized control trial. Controlled Clinical Trials 1981; 2:31-49.

37. Verhaar JAN, Walenkamp GHIM, Van Mameren H, Kester ADM, Van der Linden AJ. Local corticosteroid injection versus cyriax-type physiotherapy for tennis elbow. Journal of Bone and Joint Surgery 1996; 78-B:128-132.

38. Binder A, Hodge G, Greenwood AM, Hazleman BL, Page Thomas DP. Is therapeutic ultrasound effective in treating soft tissue lesions? British Medical Journal 1985; 290:512-514.

39. Day BH, Govindasamy N, Patnaik R. Corticosteroid injections in the treatment of tennis elbow. Practitioner 1978; 220:459-462.

40. Nevelos AB. The treatment of tennis elbow with trimcinolone acetone. Current Medical Research Opinion 1980; 6:507-509.

41. Albrecht S, Kleihues H, Cordis R, Noack W. Diagnostic and therapeutic strategies for the operative treatment of radiohumeral epicondylopathy. Archives of Orthopaedic andTrauma Surgery 1997; 116:164-72.

42. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chron Dis 1985; 38:27-36.

43. Stratford P, Levy D. Assessing valid change over time in patients with lateral epicondylitis at the elbow. Clinical Journal of Sport Medicine 1994; 4:88-91.

44. Sim J, Arnell P. Measurement validity in physical therapy research. Physical Therapy 1993; 73:102-110.

45. MacDermid J. Development of a scale for patient rating of wrist pain and disability. Journal of Hand Therapy 1996:178-183.

46. Heald SL, Riddle DL, Lamb RL. The shoulder pain and disability index: the construct validity and responsiveness of a region-specific disability measure. Physical Therapy 1997; 77:1079-1089.

47. Guyatt GH, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. Journal of Chronic Disability 1987; 40:171-178.

48. May L. The challenge of measuring change: Responsiveness of outcome measurements. Canadian Journal of Rehabilitation 1997; 10:15-24.

49. Stratford PW, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. Physical Therapy 1996; 76:1109-1123.

50. Liang M, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. Arthritis and Rheumatism 1985; 28:542-547.

51. Anderson JJ, Firschein HE, Meenan RF. Sensitivity of health status measure to short-term clinical changes in arthritis. Arthritis and Rheumatism 1989; 32:844-850.

52. Johannsen F, Gam A, Hauschild B, B. M, Jensen L. Rebox: An adjunct in physical medicine. Archives of Physical Medicine and Rehabilitation 1993; 74:438-440.

53. Huskisson EC. Vertical or horizontal visual analogue scales. Annals of Rheumatic Diseases 1979; 37:378-381.

54. Stratford P, Levy D, Gowland C. Evaluative properties used to assess patients with lateral epicondylitis at the elbow. Physiotherapy Canada 1993; 45:160-164.

55. Burton K. Grip strength in tennis elbow. British Journal of Rheumatology 1984; 23:300-301.

56. Stratford P, Norman G, McIntosh J. Generalizability of grip strength measurments in patients with tennis elbow. Physical Therapy 1989; 69:276-281.

57. Stratford P. Re: the revised pain free function index, 1997.

58. Stratford P, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. Physiotherapy Canada 1995; 47:258-263.

59. Kramer H, Theimann S. How Many Subjects?: Statistical power analysis in research. Newbury Park: Sage Publications, 1987.

60. Lees C. Interview regarding feasibility of conducting lateral epicondylitis study, 1997.

61. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin 1979; 86:420-428.

62. Stratford P. Confidence limits for your ICC. Physical Therapy 1989; 69:237-238.

63. Domholdt E. Physical therapy research: principles and applications. Philadelphia: W.B. Saunders Company, 1993.

64. Streiner DL, Norman GR. Health Measurement Scales: a practical guide to their development and use. New York: Oxford University Press, 1996.

65. Oyster CKH, W.P Llorens, L. A. Introduction to Research: a guide for the health science professional. Philadelphia: J.B. Lippincott Company, 1987.

66. Bear-Lehman J, Abreu B. Evaluating the hand: issue in reliability and validity. Physical Therapy 1989; 69:1025-1033.

67. Williams DT, BE. An empirical assessment of pain beliefs. Pain 1989; 36:351-358.

68. Solem-Bertoft EL, I. Westerberg, C. I. Pain is a major determinant of impaired performance in standardized active motor tests. Scandinavian Journal of Rehabilitation Medicine 1996; 28:71-78.

69. Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: Normative data for adults. Archives of Physical Medicine and Rehabilitation 1985; 66:69-74.

70. Abello-Banfi M, Cardiel M, Ruiz-Mercado R, Alarcon-Segou D. Quality of life in rheumatoid arthritis: validation of a Spanish version of the Arthritis Impact Measurement Scales. Journal of Rheumatology 1994; 21:1250-1255.

71. Norman GR, Stratford P, Regehr G. Methodological problems in the restrospective computation of responsiveness to change: The lesson of cronbach. Journal of Clinical Epidemiology 1997; 50:869-879.

Appendices

Appendix A Painfree Function Index

Date_____ Name _____

Today, do you or would you have any elbow discomfort at all with any of the following activities?

	discomfort	no discomfort
Dressing yourself or pulling up your slacks	_____	_____
Opening a jar or feeding yourself	_____	_____
Washing yourself or wringing out a face cloth	_____	_____
Household tasks (cleaning, lifting a chair, gardening)	_____	_____
Opening doors	_____	_____
Carrying objects with your involved hand	_____	_____
Everyday work or school activities	_____	_____
Recreation or sporting activities	_____	_____

Reprinted From: Stratford P, Levy D, Gauldie S, Levy K, Miseferi D: Extensor carpi radialis tendonitis: a validation of selected outcome measures. Physiother Can 39: 226, 1987

Appendix B Tennis Elbow Function Scale

Name Date

Today, do you or would you have any elbow discomfort at all with the following activities?

ACTIVITIES	No Discomfort	Slight Discomfort	Moderate Discomfort	Quite a bit of Discomfort	Extreme Discomfort
a. Your usual work, housework, or school activities	0	1	2	3	4
b. Your usual hobbies, sporting or recreational activities	0	1	2	3	4
c. Activities like sweeping or raking	0	1	2	3	4
d. Using tools or appliances	0	1	2	3	4
e. Dressing yourself or pulling up your pants/tights	0	1	2	3	4
f. Squeezing or gripping an object	0	1	2	3	4
g. Opening Doors with your involved limb	0	1	2	3	4
h. Carrying a small suitcase with your involved limb	0	1	2	3	4
I. Opening a jar or can	0	1	2	3	4
J. Writing or using a keyboard	0	1	2	3	4

© Stratford/Binkley 1995

Appendix C Sample Size Calculation Intraclass Correlation

Research Hypothesis
Reliability will be acceptable if ICC (1,1) > 0.75 for TEFS Test 1 and Test 2 scores.
Statistical Hypothesis
H_0: $\rho_0 = 0.75$
H_1: $\rho_1 > 0.75$
ρ_0 is the value specified in the hypothesis
ρ_1 is the estimate of the correlation that is felt to be important to find
$\rho_1 = 0.90$

ρ_1 has been selected as a desirable correlation based upon the values as stated in Portney and Watkins (1993) pg. 514

$\Delta = (\rho_1 - \rho_0) / (1 - \rho_1\rho_0)$

$\Delta = (0.90 - 0.75) / [1 - (0.90)(0.75)] = 0.462$

$\Delta = 0.45$

Based upon the formula outlined in "How Many Subjects" Kramer and Thiemann (1987) pg. 54 - 55
level of significance = 0.05 one tailed
power level = 0.8
$\Delta = 0.45$
$\nu = 28$
$n = \nu + 1 = 28 + 1$

29 subjects are required for the test retest reliability analysis, however 30 subjects will be used in the study

Appendix D Master Table for Sample Size Calculation

MASTER TABLE 5% Level, One-Tailed Test Continued

	POWER										
Δ	99	95	90	80	70	60	50	40	30	20	10
0.22	317	218	173	125	96	74	56	40	27	15	.
0.24	265	182	144	105	80	62	47	34	23	12	.
0.26	224	154	122	89	68	52	40	29	19	11	.
0.28	192	132	105	76	58	45	34	25	17	10	.
0.30	166	114	91	66	51	39	30	22	15	.	.
0.32	145	100	79	58	44	34	26	19	13	.	.
0.34	127	88	70	51	39	30	23	17	12	.	.
0.36	113	78	62	45	35	27	20	15	10	.	.
0.38	100	69	55	40	31	24	18	14	.	.	.
0.40	89	62	49	36	28	21	16	12	.	.	.
0.45	69	48	38	28	21	17	13	10	.	.	.
0.50	54	37	30	22	17	13	10	.	.	.	.
0.55	43	30	24	17	14	11	.	.	.	.	.
0.60	34	24	19	14	11	.	.	.	.	.	.
0.65	28	19	16	12	.	.	.	.	.	.	.
0.70	23	16	13	10	.	.	.	.	.	.	.
0.75	18	13	10	.	.	.	.	.	.	.	.
0.80	15	10	.	.	.	.	.	.	.	.	.
0.85	12	.	.	.	.	.	.	.	.	.	.
0.90	.	.	.	.	.	.	.	.	.	.	.

From: How Many Subjects, Kramer and Thiemann[59]

Appendix E Sample Size for Pearson Product-Moment Correlation

Based upon the formula outlined in "How Many Subjects" Kramer and Thiemann (1987) pg. 55-56

Research Hypothesis

The TEFS will be moderately correlated ($r > 0.50$) with the PFG, pain VAS, and function VAS.

Statistical Hypothesis

H_0: $r = 0.50$
H_1: $r > 0.50$

ρ_0 is the value specified in the hypothesis
ρ_1 is the estimate of the correlation that is felt to be important to find

$\rho_0 = 0.50$
$\rho_1 = 0.80$

$\Delta = (\rho_1 - \rho_0) / (1 - \rho_1\rho_0)$
$\Delta = (0.80 - 0.50) / (1 - 0.80(0.50))$
$\Delta = 0.5$

level of significance = 0.05 one tailed
power level = 0.8

$\nu = 22$
$n = \nu + 2 = 22 + 2$
$n = 24$

Appendix F Physical Therapy Clinics

Capilano Rehabilitation Centre	#100 Twin Atria Building 4999-98 Ave. T6B 2X3
Edmonton Physical Therapy Clinic Ltd.	10848 Jasper Avenue T5J 2B2
Glenora Physiotherapy Clinic	B1 10155 120 street T5K 2A2
Glen Sather Sports Medicine Clinic	E-05 Van Vliet Centre, University of Alberta T6G 2H9
Grey Nuns Community Health Centre	1100 Youville Drive West T6L 5X8
Hys Centre Physical Therapy Ltd.	#408, 11010-101 street T5H 4B8
Kinsmen Sports Centre Physical Therapy Clinic	#102 9100 Walterdale Hill T6E 2V3
Tawa Physical Therapy and Sports Injury Clinic Ltd.	#117, 3017 66 street T6K 4B2

Appendix G Inclusion Criteria Screen

Therapists please complete the following form to determine if this client is eligible for the study. Subjects are eligible to participate if they:

Yes	No	
☐	☐	complain of lateral elbow pain a minimum duration of one week
☐	☐	are able to understand spoken and written English and are able to answer the questionnaires independently
☐	☐	have access to a telephone

possess two of the following four signs:

☐	☐	lateral elbow pain with resisted wrist extension performed with the elbow extended, forearm pronated and wrist flexed
☐	☐	tenderness on palpation at the lateral epicondyle or tenoperiosteal junction
☐	☐	lateral elbow pain with resisted middle finger extension
☐	☐	lateral elbow pain with passive overpressure with forearm held in a position of full elbow extension, full wrist flexion, pronation and ulnar deviation
☐	☐	the subject has at least two of the above signs

Subjects will be excluded from the study if the answer is Yes to any of the following:

Yes	No	
☐	☐	history of fractures to the involved elbow or forearm
☐	☐	history of elbow surgery on the involved elbow
☐	☐	osteoarthritis of the involved elbow
☐	☐	active shoulder or wrist tendonitis or pathology in the involved limb
☐	☐	systemic diseases or disorders such as rheumatoid arthritis or fibromyalgia causing elbow pain

Yes	No	
☐	☐	In my opinion the subject is eligible to participate in the study
☐	☐	The subject has been provided with a subject information letter
☐	☐	The subject consents to volunteer and has agreed his/her name can be forwarded to the researcher Audrey Lowe

_____

therapist's signature

Appendix H Consent Form

Title of Project: The Test Retest Reliability, Construct Validity and Responsiveness of the Tennis Elbow Function Scale

Principal Investigator: Audrey Lowe
Master's of Science Candidate
Department of Physical Therapy
403 - 463 - 0990 (H) 403 - 450 - 7161 (W)

Co-Investigator: Dr. Michele Crites-Battie
Department of Physical Therapy
403 - 492 - 5984

Do you understand that you have been asked to be in a research study?	Yes	No
Have you read and received a copy of the attached Information Sheet?	Yes	No
Do you understand the benefits and risks involved in taking part in this research study?	Yes	No
Have you had an opportunity to ask questions and discuss this study?	Yes	No
Do you understand that you are free to refuse to participate in the study. You do not have to give a reason and it will not affect your therapy.	Yes	No
Has the issue of confidentiality been explained to you? Do you understand who will have access to your information?	Yes	No
Do you want the investigator to inform your family doctor that you are participating in this research study? If so, please provide your doctor's name: _____	Yes	No

This study was explained to me by: _____

I agree to take part in this study.

_____	_____	_____
Signature of Research Participant	Printed Name of Participant	Date
_____	_____	_____
Signature of Witness	Printed of Name Witness	Date

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

_____	_____
Signature of Investigator or Designee	Date

Appendix I Patient Information Letter

Title: The Test Retest Reliability, Construct Validity and Responsiveness of the Tennis Elbow Function Scale

Principal Investigator: Audrey Lowe
Master of Science Candidate
Department of Physical Therapy
University of Alberta
403 - 463 - 0990 (H)403 - 450 - 7161 (W)

Co- investigator Dr. Michele Crites-Battie;
Department of Physical Therapy
Faculty of Rehabilitation Medicine
University of Alberta
403 - 492 - 5984

The purpose of this study is to examine the usefulness of a questionnaire designed for people with elbow pain. You will be asked to rate the amount of discomfort or pain you have when you perform household, work and sports tasks. Results of this study will determine if this questionnaire can be used to measure peoples' response to treatments for elbow pain.

If you tell your therapist that you will volunteer for this study, your name will be forwarded to the principal investigator, Audrey Lowe. She will contact you to arrange to meet back at this clinic within the next few days. It is important that this meeting be scheduled before your second treatment for your elbow pain. At this meeting you will be asked to answer five questionnaires and you will also squeeze a handle and stop when you feel pain in your arm. The visit will take approximately one half (1/2) hour.

Within four days after your first visit with Audrey Lowe you will meet back at the clinic to answer two questionnaires. Six weeks later you will be again asked to answer five questionnaires and test grip strength. Your therapist will reexamine your arm and also answer a questionnaire. These visits will take twenty to thirty minutes. This would be considered the end of the study.

All personal information about volunteers in this study will remain confidential. Your name will not be used in any publication of the research. All data and consent forms will be securely stored in the University of Alberta, Department of Physical Therapy, Arthritis Research Lab room 179, Corbett Hall for a period of 7 years and destroyed thereafter. Only the investigators will have access to your information. Further approval from the Health Research Ethics Board will be sought prior to any further analysis of the data.

There are no risks involved in the study. You will be asked to only grip until you feel pain in your arm, thus discomfort is minimized. The researcher hopes the first two visits

will be scheduled before your second treatment with your therapist. It may be inconvenient to return to the clinic for only measurements, however this will not affect the results of your therapy. Parking expenses for the research visits will be reimbursed.

At all times you can refuse to answer any questions or withdraw from this study. The researcher will ask you for the reason why you stopped participating, however you do not have to give her an answer. Your therapy will not be affected if you do not participate or if you withdraw from the study.

Should you wish to speak to someone not involved in the study about your rights as a study participant, you may contact Dr. Herbie Rochet, Associate Dean for Graduate Studies and Research, Faculty of Rehabilitation Medicine, University of Alberta at 492-9674.

A copy of this letter and the consent form will be provided to you at the first test visit.

Thank you

I am willing to be contacted by Audrey Lowe regarding this study. I give permission to my therapist to give the researcher Audrey Lowe my name and telephone number.

_____ _____

date name

_____

telephone

I have received a copy of this information letter.

_____

Appendix J Visual Analogue Scales for Pain and Function

Date _____ Name _____

Visual Analogue Scale - Pain

Instructions place an / on the line to indicate the pain level .

Over the past 24 hours indicate your worst elbow pain level

No Pain _____ Extreme Pain

Visual Analogue Scale - Function

Place an / on the line to indicate your level of elbow function

Over the past 24 hours indicate your overall elbow function level

No Difficulty _____ Unable to use

Appendix K Subject Change Rating Scale

Change Rating Scale (Subject)

Name
Date

The purpose of this scale is for you to rate the change in your ability to function due to your **elbow pain/discomfort** since your initial treatment

☐	☐	☐
worse	same	improved

Appendix L Therapist Change Rating Scale

Change Rating Scale (Therapist)

Name
Date

The purpose of this scale is for you to rate the change in your patient's ability to function due to **elbow pain/discomfort** since initial assessment.

☐	☐	☐
worse	same	improved

Base your assessment on changes in the diagnostic criteria since the screen visit

worse	same	improved	
☐	☐	☐	pain with resisted wrist extension performed with the elbow extended
☐	☐	☐	tenderness on palpation at or close to the lateral epicondyle
☐	☐	☐	pain with resisted middle finger extension
☐	☐	☐	pain with passive overpressure with forearm held in a position of full elbow extension, full wrist flexion, pronation and ulnar deviation
☐	☐	☐	pain intensity reports
☐	☐	☐	functional ability reports

Appendix M Reliability Analysis

Test Retest Reliability

Intraclass correlation coefficient[11] pg. 511

Statistical Hypothesis
H_0: ICC = 0.75
H_1: ICC > 0.75

$$ICC(1,1) = \frac{BMS - WMS}{BMS + (k-1)\,WMS} = 0.92$$

Source of Variation	Sum of Squares	DF	Mean Square	F
Between People	3745.8387	30	124.8613	
Within People	171.0000	31	5.5161	22.6358
Between Measures	5.2258	1	5.2258	.9457
Residual	165.7742	30	5.5258	
Total	3916.8387	61	64.2105	

(n = 31) (k =2)

Confidence Interval for ICC

F value = 2.07 for (p) = .025 df (30,31)

Upper Limit F value (2.07 X 22.64)

$$R = \frac{F-1}{F + (k-1)} = 0.96$$

Lower Limit F Value (22.64 ÷2.07)
R= 0.83

Standard Error of Measurement[11]

SEM = √Ms within

SEM= √5.516 = 2.35

Appendix N Responsiveness Analysis

Magnitude of Change Scores

Research Hypothesis

TEFS scores between Test 1 and Test 2 will differ from Test 3 scores.

Statistical Hypothesis

H_0 : $\mu_1 = \mu_2 = \mu_3$

H_1 : $\mu_1 = \mu_2 > \mu_3$

$k = 3 \quad n = (31)$

$\alpha = 0.05$

Tukey's Honestly Significant Difference will be used because values significantly differ.

$| X_1 - X_2 | \geq q \sqrt{MS_e/n}$

$q = 3.49 \qquad MS_e = 20.5731 \qquad n=31$

Relative Efficiency[50]

In subjects rating themselves improved the RE of the TEFS compared to the PFF will be greater than one.

RE improved = [paired t TEFS/paired t PFF]^2 = [10.114/4.907]^2 = 4.25

Responsiveness Index

$RI^{47, 49} = \dfrac{\text{Mean Score Test 2 - Test 3}}{\text{SD of Test 1 - Test 2}}$

RI TEFS > RI PFF

Correlation of Change Scores using Pearson Correlation Coefficient

Test 1 and Test 3 change scores of the TEFS will be moderately positively correlated ($r > 0.50$) with change scores for the pain VAS, and function VAS.

Test 1 and Test 3 change scores of TEFS will be moderately negatively correlated ($r > -0.50$) with PFG.

Test 1 and Test 3 change scores of the TEFS will be strongly positively correlated ($r > 0.75$) with PFF

H_0: $r = 0.50$

H_1: $r > 0.50$

H_2: $r > -0.50$

H_3: $r > 0.75$

Where r = Pearson correlation coefficient[11] pg. 446

level of significance $p < 0.05$ one-tailed

Appendix O Letter of Approval from the Health Research Ethics Board

University of Alberta
Edmonton
Canada T6G 2G4

Faculty of Rehabilitation Medicine
Rehabilitation Research Centre
3-48 Corbett Hall
Director (403) 492-7856 Telephone (403) 492-2903
Fax (403) 492-1626

UNIVERSITY OF ALBERTA HEALTH SCIENCES FACULTIES, CAPITAL HEALTH AUTHORITY, AND CARITAS HEALTH GROUP

HEALTH RESEARCH ETHICS APPROVAL

Date: May 1998

Name(s) of Principal Investigator(s): Audrey Lowe

Organization(s): University of Alberta

Department: Graduate Studies; Department of Physical Therapy

Project Title: The Test Retest Reliability, Construct Validity, and Responsiveness to the Pain Free Function Index in Subjects with Lateral Epicondylitis.

The Health Research Ethics Board has reviewed the protocol for this project and found it to be acceptable within the limitations of human experimentation. The HREB has also reviewed and approved the patient information material and consent form.

The approval for the study as presented is valid for one year. It may be extended following completion of the yearly report form. Any proposed changes to the study must be submitted to the Health Research Ethics Board for approval.

Sharon Warren

Dr. Sharon Warren
Chair of the Health Research Ethics Board (B: Health Research)

File number: B-100598-REM

Appendix P Letter of Approval from Caritas Research Steering Committee

CARITAS HEALTH GROUP

16940 - 87 Avenue Edmonton, Alberta T5R 4H5 Tel. (403) 484-8811 Fax. (403) 930-5774

June 4, 1998

Ms. Audrey Lowe
Department of Physical Therapy
Grey Nuns Community Hospital & Health Centre
Caritas Health Group
Edmonton, Alberta

Dear Ms. Lowe:

RE: The Test-Retest Reliability, Construct Validity and Responsiveness of the Pain Free Function Index

Thank you for submitting this study to the Caritas Research Steering Committee. This study has received ethical approval from the Health Research Ethics Board, which is accepted by Caritas. I am pleased to inform you that the Caritas Research Steering Committee has granted Administrative Approval for this study at the Grey Nuns Hospital.

We would appreciate a report to our Committee on an annual basis and at the completion of the study.

All financial arrangements must be submitted and approved by Ms. Vikki Newman, Coordinator, Treasury and Risk Management, Finance Department, Caritas Health Group.

Your request for a grant from the Caritas Research Steering Committee will be considered in the near future, and Peggy Morton will contact you in this regard.

If you have any questions, please do not hesitate to contact me. You can page me at the Grey Nuns Hospital or leave a message with the Committee secretary, Peggy Morton at 930-5924 or fax 930-5961.

Yours truly,

G.F. MacDonald, M.D., FRCP(C)
Chairperson, Caritas Research Steering Committee

Members: Edmonton General
Misericordia Community Hospital and Health Centre

Appendix B

Revised Pain Free Function Index

Today, do you or would you have any elbow discomfort at all with the following activities?

(circle one number on each line)

ACTIVITIES	No Discomfort	Slight Discomfort	Moderate Discomfort	Quite a bit of Discomfort	Extreme Discomfort
a. Your usual work, housework, or school activities	0	1	2	3	4
b. Your usual hobbies, sporting or recreational activities	0	1	2	3	4
c. Activities like sweeping or raking	0	1	2	3	4
d. Using tools or appliances	0	1	2	3	4
e. Dressing yourself or pulling up your pants/tights	0	1	2	3	4
f. Squeezing or gripping an object	0	1	2	3	4
g. Opening doors with your involved limb	0	1	2	3	4
h. Carrying a small suitcase with your involved limb	0	1	2	3	4
i. Opening a jar or can	0	1	2	3	4
j. Writing or using a keyboard	0	1	2	3	4

* Permission to Reprint Audrey Lowe

©Stratford/Binkley 1995

Date: Sat, 17 Jan 1998 08:17:54 -0500
From: Paul Stratford <straforfhs.csu.McMaster.CA>
Reply-To: stratforfhs.csu.McMaster.CA
: K Lowe <loweegpu.srv.ualberta.ca>
Subject: Re: The Pain Free Function Index (fwd)

Hi Audrey
You are welcome to modify the scale as long as you state it. We want patients and clinicians to have the best tool that is available (not our tool). With respect to the numbers: You want it easy for clinicians to score. Omitting numbers can be dangerous. When scores are attached, some clinicians may invert the intended scoring scheme. Thus, in some cases highs scores will represent positive function and in other cases low scores will represent positive function. Based, on the literature, I'm not sure (other than using negative score values) that including numbers is a problem. Indeed, the SF-36 folks include numbers. As a matter of interest only, can you provide my with literature supporting this position.
I am happy to review drafts or questions that you may have along the way. Please feel free to contact me at any time. Sorry to take so long getting back to you.

Regards
Paul

Date: Thu, 28 Aug 1997 07:51:02 -0400
From: Paul Stratford <strafor@fhs.csu.McMaster.CA>
Reply-To: stratfor@fhs.csu.McMaster.CA
To: K Lowe <lowe@gpu.srv.ualberta.ca>
bject: Re: The revised Pain Free Function Index

Hello Audrey
Items were generated by using a patient specific functional scale. Here patients identified items they were having difficulty with. Following this, item analysis was performed. You have our blessing to do a reliability, validity, and sensitivity to change study using the instrument. I would be interested in reviewing the study design, but this is not a requirement to you using it. I wish you the best, and please keep us posted.

Regards
Paul

Ortho journal -7